Risk Analysis, Vol. 13, No. 2, 1993

Upper Confidence Limits on Excess Risk for Quantitative Responses

Ralph L. Kodell[1,3] and Ronnie W. West[1,2]

Received March 12, 1992; revised August 10, 1992

The definition and observation of clear-cut adverse health effects for continuous (quantitative) responses, such as altered body weights or organ weights, are difficult propositions. Thus, methods of risk assessment commonly used for binary (quantal) toxic responses such as cancer are not directly applicable. In this paper, two methods for calculating upper confidence limits on excess risk for quantitative toxic effects are proposed, based on a particular definition of an adverse quantitative response. The methods are illustrated with data from a dose–response study, and their performance is evaluated with a Monte Carlo simulation study.

KEY WORDS: Additional risk; benchmark dose; likelihood ratio; noncentral *t*.

1. INTRODUCTION

In the case of binary (quantal) toxic responses such as cancer and frank birth defects, the definition of an adverse health effect is self-evident and the toxic event can be observed on individual subjects. By contrast, a clear-cut adverse effect for a continuous (quantitative) response such as altered blood concentration of a toxicant, altered body weight, or altered organ weight is difficult both to define and to observe unequivocally. Perhaps this partly explains why the widely accepted methods of risk assessment for quantal toxic endpoints have not carried over to quantitative responses; whereas the risk of adverse quantal responses can be modeled easily in terms of the probability of occurrence of such effects, characterizing the risk of quantitative responses in terms of probability of occurrence does not follow as naturally.

[1] Biometry and Risk Assessment, HFT 20, National Center for Toxicological Research, NCTR Drive, Jefferson, Arkansas 72079.
[2] Present address: Department of Statistics, Rice University, Houston, Texas 77251.
[3] To whom all correspondence should be addressed.

There have been efforts to develop methods of risk assessment for continuous quantitative responses.[1–6] Except for the approach of Chen and Gaylor,[6] methods that characterize the risk of quantitative effects on a probability scale are restricted to point estimates of the risk of such effects. On the other hand, Chen and Gaylor[6] estimate statistical upper confidence limits on that risk. Crump[2] also employs upper confidence limits for continuous effects, but does not characterize the risk in terms of probability of occurrence. Presently accepted methods of risk assessment for carcinogenic effects utilize statistical upper confidence limits on the probability of such effects. These methods are used to obtain upper bounds on the true excess risk above background risk. It is considered inadvisable to rely solely on model-based point predictions of risk for setting acceptable levels of exposure to carcinogenic substances, due to the wide variability in the extrapolated predictions of various models that fit well in the data range.[7,8] The present paper presents two methods for obtaining statistical upper confidence limits on excess risk for continuous quantitative toxic responses.

0272-4332/93/0400-0177$07.00/1

2. PROBABILITY DISTRIBUTION AND DOSE–RESPONSE MODEL

Suppose that there are g dose groups ($i=1,2,\ldots,g$) in a dose–response experiment, with d_i denoting the dose level and n_i the sample size for the i^{th} group, and $N=\Sigma n_i$ denoting the total size of the experiment. Let $X(d)$ represent the quantitative response variable of interest, with $x_j(d_i)$ denoting the observed value of $X(d)$ on the j^{th} experimental unit in group i ($j=1,2,\ldots,n_i$). For most continuous-type responses encountered in toxicology, either a normal or a lognormal probability distribution will adequately describe the data.[9,10] For the present procedure, a normal distribution is assumed for the response variable $X(d)$ at each dose level. The likelihood function is

$$L = \prod_{i=1}^{g} \prod_{j=1}^{n_i} \frac{1}{\sigma\sqrt{(2\pi)}} \exp\{-1/2[x_j(d_i) - \mu(d_i)]^2/\sigma^2\}$$

where $\mu(d_i)$ is the mean of the distribution of $X(d_i)$, and σ, the standard deviation, is a nuisance parameter that is assumed to be the same for all dose levels. Let $\mu(d)$ be expressed by a polynomial in d of degree at most $g-1$. Often a second-degree polynomial should suffice, in which case

$$\mu(d) = \beta_0 + \beta_1 d + \beta_2 d^2$$

Once an appropriate error distribution (normal) and a suitable dose–response model (second-degree polynomial) have been selected, the next step is to define an abnormal effect in terms of $X(d)$.

3. PROBABILITY OF AN ABNORMAL EFFECT

In the same manner as that of the Safe Drinking Water Committee of the National Research Council[1] and of Gaylor and Slikker,[3] an abnormal adverse effect is defined to be a value of $X(d)$ whose magnitude corresponds to values in the tail area of the control (background) distribution [i.e., the distribution of $X(0)$]. Consequently, abnormal values are those that are far from the control mean and which occur with low probability in unexposed subjects. Values in either the upper or lower tail may be considered adversely abnormal, but only values in the lower tail will be considered here. Thus, an abnormal response is defined to be one smaller than $\mu(0)-k\sigma$, where k is appropriately chosen to yield a specific low percentage point. For example, $k=2.33$ gives a tail probability of 0.01, while $k=3.62$ gives 0.0001. In effect, this definition of an abnormal effect dichotomizes the continuous response variables; however, the dichotomized response generally is not explicitly observable, since $\mu(0)$ generally is unknown and must be estimated.

Under the assumed normal distribution, the probability of an adverse effect at dose level d is given by

$$\begin{aligned}&\text{Prob}[X(d) < \mu(0) - k\sigma]\\ &= \text{Prob}\{[X(d) - \mu(d)]/\sigma \le [\mu(0) - \mu(d) - k\sigma]/\sigma\}\\ &= \Phi\{[\mu(0) - \mu(d)]/\sigma - k\}\end{aligned}$$

where Φ denotes the standard normal cumulative distribution function. In this paper, the excess risk above background is characterized by "additional risk," which is represented by

$$\begin{aligned}\pi(d) &= \text{Prob}[X(d) < \mu(0) - k\sigma]\\ &\quad - \text{Prob}[X(0) < \mu(0) - k\sigma]\\ &= \Phi\{[\mu(0) - \mu(d)]/\sigma - k\} - \Phi\{-k\}\\ &= \Phi\{[-\beta_1 d - \beta_2 d^2]/\sigma - k\} - \Phi\{-k\}\end{aligned}$$

4. UPPER CONFIDENCE LIMITS ON EXCESS (ADDITIONAL) RISK

Under the presently assumed normal distribution and linear (in the parameters) model, either maximum-likelihood estimation or least-squares regression may be used to estimate parameters, since these methods will give equivalent results. Point estimates of additional risk above background for selected dose levels may be calculated by substituting estimates of β_1, β_2, and σ into the above expression for excess risk.

Two procedures for obtaining a one-sided $100(1-\alpha)\%$ statistical upper confidence limit on additional risk, $\pi(d_0)$, at a given dose level, d_0, will be proposed here. The first is based on the chi-square distribution of the likelihood ratio statistic and the second on the noncentral-t distribution of a function of the maximum-likelihood estimator of $\pi(d_0)$.

4.1. The Likelihood Ratio/Chi-Square Approach

The first procedure uses the distribution of the likelihood ratio statistic.[11] Likelihood ratio-based confidence limits often are used in risk assessment for carcinogenic effects.[12] With excess risk characterized by "additional risk," the upper confidence limit is defined to be the maximum value of $\pi(d_0)$ that satisfies

$$\pi(d_0) = \Phi\{[\mu(0) - \mu(d_0)]/\sigma - k\} - \Phi\{-k\}$$

and

$$2[\ln L(\hat{\Theta}) - \ln L(\tilde{\Theta})] = \chi^2_{1,2\alpha}$$

where Θ denotes the vector of parameters in the likelihood function, $\ln L(\hat{\Theta})$ is the maximum unconstrained log likelihood, and $\ln L(\tilde{\Theta})$ is the maximum log likelihood constrained by the excess-risk constraint.

An algorithm for calculating an upper confidence limit for $\pi(d_0)$ using the likelihood approach is now described. Calculate the maximum log likelihood, $\ln L(\hat{\Theta})$. For a fixed dose, $d = d_0$, select a starting value for $\pi(d_0)$, say $\pi_u(d_0) = \hat{\pi}(d_0) + \epsilon$, for some small $\epsilon > 0$, where $\hat{\pi}(d_0)$ is the unrestricted maximum-likelihood estimate of additional risk at d_0. Maximize $\ln L(\tilde{\Theta})$, subject to the constraint on excess risk (discussed below). Calculate $2[\ln L(\hat{\Theta}) - \ln L(\tilde{\Theta})]$. If $2[\ln L(\hat{\Theta}) - \ln L(\tilde{\Theta}) < \chi^2_{1,2\alpha}$, then increase the value of $\pi_u(d_0)$ by a small amount. Repeat this process until $2[\ln L(\hat{\Theta}) - \ln L(\tilde{\Theta})] = \chi^2_{1,2\alpha}$. The resulting value of $\pi_u(d_0)$ is an asymptotic $100(1-\alpha)\%$ confidence limit on excess risk at d_0. This algorithm may also be used to calculate a lower confidence limit on the dose, d_0, corresponding to a fixed excess risk, $\pi(d_0)$. The procedure works the same way as the above procedure, except that one iterates on dose rather than on excess risk.

The excess-risk constraint requires that for a given trial value of $\pi_u(d_0)$, the parameters β_1, β_2, and σ can take on only values consistent with $\pi_u(d_0) = \phi\{[\mu(0) - \mu(d_0)]/\sigma - k\} - \phi\{-k\}$. Since only a single constraint is imposed on the parameter space, the χ^2 has 1 degree of freedom. Incorporation of the excess-risk constraint for computational purposes may be accomplished as follows. Solving the excess-risk constraint for the parameter β_2 yields

$$\beta_2 = -\sigma\{k + \Phi^{-1}[\pi(d_0) + \Phi(-k)]\}/d_0^2 - \beta_1/d_0$$

Substituting this constraint back into the original model gives

$$\mu(d) = \beta_0 + \beta_1(d - d^2/d_0) - \sigma\{k + \Phi^{-1}[\pi(d_0) + \Phi(-k)]\}d^2/d_0^2$$

The maximum-likelihood solution for this restricted model is available from the authors. The solution is specific to a second-degree polynomial for mean response, whereas $\mu(d)$ can actually have degree up to $g-1$. Higher degrees would require a more general solution, but it is felt that degree = 2 will be sufficient to describe many dose–response data sets encountered in toxicity studies.

4.2. The Maximum-Likelihood Estimator/ Noncentral-*T* Approach

The second approach to obtaining an upper confidence limit on the additional risk uses the noncentral t distribution of the maximum-likelihood estimator of $[\mu(0) - \mu(d_0)]/\sigma = k + \Phi^{-1}\{\pi(d_0) + \Phi(-k)\}$. Let $\hat{\mu}(d) = \hat{\beta}_0 + \hat{\beta}_1 d + \hat{\beta}_2 d^2$ be the ordinary least-squares estimator of $\mu(d)$. The distribution of $\hat{\mu}(0) - \hat{\mu}(d_0)$ is normal with mean $\mu(0) - \mu(d_0)$ and variance $K(d_0)\sigma^2$, where $K(d_0)$ is a known constant.[13] [The expression for $K(d_0)$ is available from the authors.] Now the distribution of $[\hat{\mu}(0) - \hat{\mu}(d_0)]/[\hat{\sigma}\sqrt{K(d_0)}]$ is $T_{N-3}(\delta)$, that is, noncentral t with $N-3$ degrees of freedom and noncentrality parameter $\delta = [\mu(0) - \mu(d_0)]/[\sigma\sqrt{K(d_0)}]$,[14] where $\hat{\sigma}^2$ is the residual variance estimator of σ^2 based on N-3 degrees of freedom. As N gets large, the distribution of

$$Z = \{T_{N-3}(\delta) - E[T_{N-3}(\delta)]\}\sqrt{V[T_{N-3}(\delta)]}$$

is approximately normal with zero mean and unit variance, where $E[T_{N-3}(\delta)]$ and $V[T_{N-3}(\delta)]$ are given in Ref. 14 (pp. 203–204) as

$$E[T_{N-3}(\delta)] = \delta[(N-3)/2]^{1/2}\Gamma[(N-4)/2]/\Gamma[(N-3)/2]$$

and

$$V[T_{N-3}(\delta)] = (1 + \delta)^2[(N-3)/(N-5)] - \{E[T_{N-3}(\delta)]\}^2$$

where Γ denotes the gamma function.

To calculate an upper confidence limit on $\pi(d_0)$, the standardized statistic Z is used. Since $P(Z \geq z_{1-\alpha}) \approx 1-\alpha$, where $z_{1-\alpha}$ is the $100\alpha^{\text{th}}$ percentage point of the standard normal distribution, then an approximate upper confidence limit for δ may be found by equating Z to $z_{1-\alpha}$. An algorithm for doing this is as follows. Select a starting value for δ, say $\delta_u = \hat{\delta} + \epsilon$, where ϵ is a small positive number and $\hat{\delta} = [\hat{\mu}(0) - \hat{\mu}(d_0)]/[\hat{\sigma}\sqrt{K(d_0)}]$. Calculate Z with the given value of δ_u. If $Z > z_{1-\alpha}$, then increase the value of δ_u by a small amount. Repeat this process until $Z = z_{1-\alpha}$. Using the resulting value of δ_u, the asymptotic $100(1-\alpha)\%$ confidence limit on $\pi(d_0)$ is given by $\Phi[\delta_u\sqrt{K(d_0)} - k] - \Phi(-k)$. Just as with the likelihood ratio/chi-square (LR/CHISQ) approach, this maximum-likelihood estimator/noncentral-t (MLE/NCT) method may be used to solve iteratively for a lower confidence limit on the dose, d_0, corresponding to a fixed excess risk, $\pi(d_0)$.

5. EXAMPLE

Data from a 14-day range finding study on aconiazide conducted at the National Center for Toxicological Research (NCTR) are used to illustrate the procedures described in Sections 4.1 and 4.2. The purpose of the study was to determine via a number of toxic manifestations the appropriate dose levels of aconiazide to be used in a 6-month, repeat-dose toxicity study. Complete details of the study are provided in the NCTR Technical Report for Experiment No. 6584.02.[15] For this example, only body weight changes in female rats are analyzed.

The data and best fitting quadratic model obtained by maximum-likelihood estimation are shown in Fig. 1. (A tabular representation of the data is available from the authors.) The quadratic model appears to provide a good description of the data. Both confidence limit procedures were used to solve the reverse of the upper confidence limit problem (i.e., to determine a lower 95% confidence limit on the dose corresponding to an additional risk of 0.01) with background risk equal to $\Phi(-3)=0.0013$ (i.e., $k=3$). Such a dose level has been termed a benchmark dose (BD) for purposes of risk assessment.[2,16] For the aconiazide data, each method gave a value of 178 mg/kg/day as the benchmark dose corresponding to 1% additional risk above background. Thus, with 95% confidence, 178 mg/kg/day in a 14-day study represents a lower limit on the dose of aconiazide at which female rats would have 1% additional risk of experiencing weight loss more extreme than 3 standard deviations below the mean of the control distribution of body weight charges.

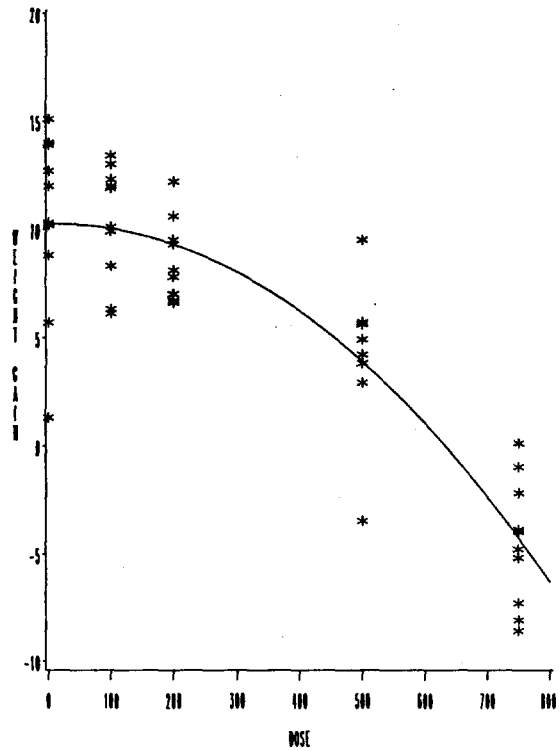

Fig. 1. Body weight gain or loss in grams for female Fischer 344 rats treated by gavage with aconiazide at various doses in mg/kg body weight daily for 14 days, along with maximum-likelihood estimated quadratic dose–response model.

6. MONTE CARLO SIMULATION STUDY

In order to verify the nominal coverage levels of the two proposed methods for obtaining upper confidence limits on additional risk when all assumptions are met, a limited Monte Carlo simulation study was conducted in which simulated coverage probabilities were compared to the nominal level. Both GLIM[17] and S-PLUS[18] software were used to generate the data and compile the results. An experiment with five dose groups including a control was assumed (dose levels: $d_i=i$, $i=0$, 1, 2, 3, 4). Two cases were considered. For case 1, each dose group was simulated to have 10 animals ($n_i=10$, $N=50$); while for case 2, each dose group was simulated to have 20 animals ($n_i=20$, $N=100$). For each of the two cases, 1000 data sets wre simulated using the model

$$X_{ij} = 3 - d_i - 0.1d_i^2 + \epsilon_{ij}$$

where ϵ_{ij} was generated as $N(0,1)$ for $i=0,\ldots,4$ and $j=1,\ldots,n_i$ ($n_i=10$ or 20). For dose groups 1, 2, and 3, one-sided 95% upper confidence limits on additional risk (defined with $k=3$) were calculated for each of the 1000 data sets for each sample size, and these were compared to the true additional risk at each of the three dose levels. For a given dose level, the estimated coverage probability was calculated as the proportion of the 1000 simulations for which the resulting upper confidence limit equaled or exceeded the true additional risk. The results are given in Table I for both confidence limit procedures. Also included in Table I are coverage probabilities for lower confidence limits on the 1% benchmark dose.

7. DISCUSSION

From the simulation results reported in Table I, it can be seen that both methods for calculating confidence limits provided good coverage with respect to the nom-

Table I. Simulated Coverage Probabilities of Upper Confidence Limits on Additional Risk for Fixed Dose and Lower Confidence Limits on Dose for Fixed Additional Risk for Nominal 95% Confidence Level

			Method[a]					
			LR/CHISQ			MLE/NCT		
Sample size	Dose level	True risk	Avg.[b] UCL	Var.[c] UCL	Cov.[d] Prob.	Avg. UCL	Var. UCL	Cov. Prob.
$n_i=10$	1	0.027	0.097	0.0037	0.957	0.093	0.0035	0.959
($N=50$)	2	0.273	0.595	0.0329	0.961	0.592	0.0298	0.968
	3	0.815	0.953	0.0029	0.972	0.951	0.0033	0.963
$n_i=20$	1	0.027	0.067	0.0008	0.956	0.065	0.0008	0.946
($N=100$)	2	0.273	0.495	0.0181	0.961	0.492	0.0172	0.949
	3	0.815	0.932	0.0030	0.954	0.929	0.0031	0.956
Sample size	Dose level	True risk	Avg.[b] LCL	Var.[c] LCL	Cov.[d] Prob.	Avg. LCL	Var. LCL	Cov. Prob.
$n_i=10$	0.676	0.01	0.460	0.0128	0.956	0.464	0.0127	0.948
$n_i=20$	0.676	0.01	0.505	0.0079	0.958	0.509	0.0078	0.945

[a] LR/CHISQ is the likelihood ratio/chi-square method and MLE/NCT is the maximum-likelihood estimator/noncentral-t method.
[b] Average value of upper confidence limits (UCL) or lower confidence limits (LCL) based on 1000 simulations.
[c] Sample variance of the UCL or LCL based on 1000 simulations.
[d] Proportion of 1000 simulations for which the calculated UCL or LCL covered the true value.

inal 95% level. Both methods had coverage probabilities closer to nominal for the larger-sized experiment, and were slightly conservative for the smaller-sized experiment. Based on this simulation study, a slight edge might be given to the MLE/NCT approach over the LR/CHISQ approach in terms of average size of the confidence limits. Conversely, a slight edge might be given to the LR/CHISQ approach over the MLE/NCT approach in terms of maintaining nominal coverage. However, the results are too close to permit a definitive judgment to be made.

Stiteler and Durkin[14] discussed the basic approach used in this paper for defining an abnormal effect, but they expressed preference for a different approach. The present definition needs to be evaluated further with respect to model misspecifications, alternative probability distributions, and, in particular, unequal variances among the dose groups. In special cases of unequal variances, it might be possible to apply a variance-stabilizing transformation prior to analysis. In general, however, the definition of additional risk will not be preserved under arbitrary transformations. Nonetheless, the definition used here does provide an approach to assessing risk for quantitative toxic responses in the familiar and customary framework of quantitative risk assessment, and both confidence limit procedures derived on the basis of this definition perform well at least when the underlying assumptions are satisfied.

REFERENCES

1. National Research Council, "Drinking Water and Health, Vol. 3" (Report of the Safe Drinking Water Committee, Board on Toxicology and Environmental Health Hazards, Assembly of Life Sciences, Washington, D.C., National Academy Press, 1980).
2. K. S. Crump, "A New Method for Determining Allowable Daily Intake," *Fundamental and Applied Toxicology* 4, 854–871 (1984).
3. D. W. Gaylor, and W. L. Slikker, "Risk Assessment for Neurotoxic Effects," *NeuroToxicology* 11, 211–218 (1990).
4. W. M. Stiteler, and P. R. Durkin, "Some Statistical Issues Relating to the Characterization of Risk for Toxic Chemicals" (Proceedings of the Workshop on Superfund Hazardous Waste: Statistical Issues in Characterizing a Site: Protocols, Tools, and Research Needs, Arlington, Virginia, February 21–22, 1990).
5. J. R. Glowa, "Dose-Effect Approaches to Risk Assessment," *Neuroscience and Biobehavioral Reviews* 15,153–158 (1991).
6. J. J. Chen, and D. W. Gaylor, "Risk Assessment for Quantitative Response" (Technical Report for Experiment No. 6668, National Center for Toxicological Research, Jefferson, Arkansas, 1992).
7. FDA Advisory Committee on Protocols for Safety Evaluation, "Protocols for Safety Evaluation: Panel on Carcinogenesis Report on Cancer Testing in the Safety Evaluation of Food Additives and Pesticides," *Toxicology and Applied Pharmacology* 20, 419–438 (1971).
8. Office of Science and Technology Policy, "Chemical Carcinogens: A Review of the Science and Its Associated Principles"

(U.S. Interagency Staff Group on Chemical Carcinogenesis, *Federal Register* **50**, 10371–10442, 1985).

9. D. S. Salsburg, *Statistics for Toxicologists* (New York, Marcel Dekker, 1986).
10. B. D. Litt, K. E. Boyle, and L. E. Myers, "Statistical Analysis of Subchronic Toxicity Studies, in D. Krewski and C. Franklin (eds.), *Statistics in Toxicology* (New York, Gordon and Breach, 1991).
11. D. R. Cox and D. V. Hinkley, *Theoretical Statistics* (London, Chapman and Hall, 1974).
12. K. S. Crump and R. B. Howe, "A Review of Methods for Calculating Confidence Limits in Low Dose Extrapolation, in D. B. Clayson, D. R. Krewski, and I. C. Munroe (eds.), *Toxicological Risk Assessment, Vol. 1: Biological and Statistical Criteria* (Boca Raton, CRC Press, 1985), pp. 187–203.
13. P. Bickel and K. Doksum, *Mathematical Statistics* (Oakland, Holden-Day, 1977).
14. N. L. Johnson and S. Kotz, *Continuous Univariate Distributions-2* (New York, Houghton Mifflin, 1970).
15. National Center for Toxicological Research, "Fourteen-Day, Repeat-Dose, Dose-ranging Toxicity Evaluation of Aconiazide in Male and Female Fisher 344 Rats" (Technical report for Experiment No. 6584.02, Division of Biochemical Toxicology, NCTR, Jefferson, Arkansas, 1990).
16. J. J. Chen and R. L. Kodell, "Quantitative Risk Assessment for Teratological Effects," *Journal of the American Statistical Association* **84**, 966–971 (1989).
17. GLIM, "The Generalized Linear Interactive Modelling System" (Release 3.77. Numerical Algorithms Group, Inc., Downers Grove, Illinois, 1987).
18. S-PLUS (Statistical Sciences, Inc., Seattle, Washington, 1990).